R Workshop

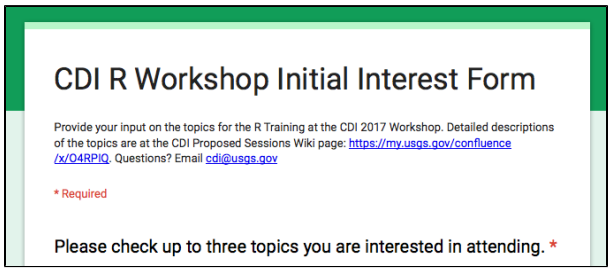
Brief summary of the proposed session:

USGS-R (https://owi.usgs.gov/R/) runs R training workshops for a range of skill levels from beginner to advanced. We are working with the USGS-R team to bring trainings to the CDI meeting.

Point of Contact: currently, Leslie Hsu and Madison Langseth at cdi@usgs.gov, but we will update with R instructors as this evolves

Envisioned format: Full or half day hands-on workshops on the Friday of the meeting, shorter 1.5 hour breakouts may occur earlier in the week.

****Mark your preferred topics on this form to help with the workshop planning. ****


CDI R Workshop Initial Interest Form

Provide your input on the topics for the R Training at the CDI 2017 Workshop. Detailed descriptions of the topics are at the CDI Proposed Sessions Wiki page: https://my.usgs.gov/confluence/x/O4RPIQ. Questions? Email cdi@usgs.gov

* Required

Please check up to three topics you are interested in attending. *

Possible Topics (and see Table below for more information)

1. **GitHub with R:** Learn the basics of version control and the collaborative online software repository GitHub
2. **HTC/HPC:** Learn how to use high throughput/high performance computing with your code
3. **Intro to R:** Learn the basics of scientific computing in R
4. **SBtools:** Learn scripted ways to send data to and from ScienceBase
5. **dataRetrieval:** Learn about: dataRetrieval, an R-package designed to simplify data access to USGS and EPA water data.
6. **USGS Packages a la carte**: Learn about R tool development within the USGS
7. **Intermediate R** : data cleaning, dplyr, tidyr, data.tables, functions: Take your R skills to the next level: Learn powerful new packages that make it faster to write and run code

	Potential course length			Brief description / learning objectives	Existing course material (if available)	Citation or docum
	1.5hr	3hr	1d			
Introduction to R: Learn the basics of scientific computing in the R programming environment			x	Learn the basics of scientific computing in R, including how to • Import data into R from local files or web sources • Reformat and 'clean' data to prepare for analysis • Analyze and visualize data • Repeat and reproduce	Yes: https://owi.usgs.gov/R/training-curriculum/intro-curriculum/	

Intermediate R		x	x	Take your R skills to the next level: Learn powerful new packages that make it faster to write and run code • concisely implement the split-apply-combine paradigm with dplyr and tidyr • read files blazingly fast with data.tables • increase code reusability by writing your own functions	Could be pulled from https://owi.usgs.gov/R/training-curriculum/intro-curriculum/ and expanded to ~1 day	https://www.rstudio.com/wp-content/uplo , http://adv-r.had.co.nz/Functions.html
USGS Packages a la carte		x	x	An overview of USGS R packages for accessing and analyzing environmental data: • dataRetrieval • geoknife • sbtools • EGRET • And more…!	Yes: http://usgs-r.github.io/a-la-carte/, but would need to be updated/refined	
SBtools	x	x		Learn scripted ways to send data to and from ScienceBase • query ScienceBase for existing publications and datasets • retrieve and visualize data from ScienceBase • create your own ScienceBase items and attach your own files	No	https://cran.r-project.org/web/packages/s org/archive/2016-1/winslow-chamberlain
dataRetrieval	x	x		Learn about dataRetrieval, an R-package designed to simplify data access to USGS and EPA water data. • Discover • Retrieve • Explore	https://owi.usgs.gov/R/dataRetrieval.html#1	https://cran.r-project.org/web/packages/d
Collaborative software development and version control using Git, GitHub, and R		x	x	Learn the basics of version control and collaborative online software repository Github • Get git set up on local machine • Establish github account and create and merge pull requests • ...	NA, but see preliminary documentation here: https://github.com/USGS-R/wateRuse/issues/15	

High throughput computing (HTC) with R		x	x	Learn how to use high throughput/high performance computing with your code • Practice transitioning between serial and parallel code implementations • Learn about USGS cluster computing resources and try them out	NA, but see http://usgs-r.github.io/a-la-carte/htcondor.pptx	